

10/589594

IAP11 Rec'd PCT/PTO 15 AUG 2006

SEQUENCE LISTING

<110> Nakamura, Yusuke
Furukawa, Yoichi

<120> METHOD FOR DIAGNOSING COLORECTAL CANCERS

<130> 082368-008900US

<150> PCT/JP2004/002145

<151> 2004-02-24

<160> 23

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2624

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (294)...(1688)

<400> 1

```

ggcacgaggg gccgacgcga gcgcgcgcgt tcgcttcagc tgctagctgg cccaagggag 60
gcgaccgcgg aggggtggcga ggggcggcca ggacccgcag ccccggggcc gggccgggtcc 120
ggaccgccag ggagggcagg tcagtgggca gatcgcgctcc gcgggattca atctctgccc 180
gctctgataa cagtcctttt ccctggcgct cacttcgtgc ctggcacccg gctgggcgcc 240
tcaagaccgt tgtctcttcg atcgcttctt tggacttggc gaccatttca gag atg 296
                                         Met
                                         1

```

```

tct tcc aga agt acc aaa gat tta att aaa agt aag tgg gga tcg aag 344
Ser Ser Arg Ser Thr Lys Asp Leu Ile Lys Ser Lys Trp Gly Ser Lys
                    5                      10                      15

```

```

cct agt aac tcc aaa tcc gaa act aca tta gaa aaa tta aag gga gaa 392
Pro Ser Asn Ser Lys Ser Glu Thr Thr Leu Glu Lys Leu Lys Gly Glu
                20                      25                      30

```

```

att gca cac tta aag aca tca gtg gat gaa atc aca agt ggg aaa gga 440
Ile Ala His Leu Lys Thr Ser Val Asp Glu Ile Thr Ser Gly Lys Gly
                35                      40                      45

```

```

aag ctg act gat aaa gag aga cac aga ctt ttg gag aaa att cga gtc 488
Lys Leu Thr Asp Lys Glu Arg His Arg Leu Leu Glu Lys Ile Arg Val
                50                      55                      60                      65

```

```

ctt gag gct gag aag gag aag aat gct tat caa ctc aca gag aag gac 536
Leu Glu Ala Glu Lys Glu Lys Asn Ala Tyr Gln Leu Thr Glu Lys Asp
                    70                      75                      80

```

```

aaa gaa ata cag cga ctg aga gac caa ctg aag gcc aga tat agt act 584
Lys Glu Ile Gln Arg Leu Arg Asp Gln Leu Lys Ala Arg Tyr Ser Thr
                    85                      90                      95

```

```

acc gca ttg ctt gaa cag ctg gaa gag aca acg aga gaa gga gaa agg 632

```

Thr	Ala	Leu	Leu	Glu	Gln	Leu	Glu	Glu	Thr	Thr	Arg	Glu	Gly	Glu	Arg	
		100					105					110				
agg	gag	cag	gtg	ttg	aaa	gcc	tta	tct	gaa	gag	aaa	gac	gta	ttg	aaa	680
Arg	Glu	Gln	Val	Leu	Lys	Ala	Leu	Ser	Glu	Glu	Lys	Asp	Val	Leu	Lys	
		115				120					125					
caa	cag	ttg	tct	gct	gca	acc	tca	cga	att	gct	gaa	ctt	gaa	agc	aaa	728
Gln	Gln	Leu	Ser	Ala	Ala	Thr	Ser	Arg	Ile	Ala	Glu	Leu	Glu	Ser	Lys	
		130				135				140					145	
acc	aat	aca	ctc	cgt	tta	tca	cag	act	gtg	gct	cca	aac	tgc	ttc	aac	776
Thr	Asn	Thr	Leu	Arg	Leu	Ser	Gln	Thr	Val	Ala	Pro	Asn	Cys	Phe	Asn	
				150					155					160		
tca	tca	ata	aat	aat	att	cat	gaa	atg	gaa	ata	cag	ctg	aaa	gat	gct	824
Ser	Ser	Ile	Asn	Asn	Ile	His	Glu	Met	Glu	Ile	Gln	Leu	Lys	Asp	Ala	
			165					170					175			
ctg	gag	aaa	aat	cag	cag	tgg	ctc	gtg	tat	gat	cag	cag	cgg	gaa	gtc	872
Leu	Glu	Lys	Asn	Gln	Gln	Trp	Leu	Val	Tyr	Asp	Gln	Gln	Arg	Glu	Val	
		180					185					190				
tat	gta	aaa	gga	ctt	tta	gca	aag	atc	ttt	gag	ttg	gaa	aag	aaa	acg	920
Tyr	Val	Lys	Gly	Leu	Leu	Ala	Lys	Ile	Phe	Glu	Leu	Glu	Lys	Lys	Thr	
	195					200					205					
gaa	aca	gct	gct	cat	tca	ctc	cca	cag	cag	aca	aaa	aag	cct	gaa	tca	968
Glu	Thr	Ala	Ala	His	Ser	Leu	Pro	Gln	Gln	Thr	Lys	Lys	Pro	Glu	Ser	
						215				220					225	
gaa	ggt	tat	ctt	caa	gaa	gag	aag	cag	aaa	tgt	tac	aac	gat	ctc	ttg	1016
Glu	Gly	Tyr	Leu	Gln	Glu	Glu	Lys	Gln	Lys	Cys	Tyr	Asn	Asp	Leu	Leu	
				230					235					240		
gca	agt	gca	aaa	aaa	gat	ctt	gag	gtt	gaa	cga	caa	acc	ata	act	cag	1064
Ala	Ser	Ala	Lys	Lys	Asp	Leu	Glu	Val	Glu	Arg	Gln	Thr	Ile	Thr	Gln	
			245					250					255			
ctg	agt	ttt	gaa	ctg	agt	gaa	ttt	cga	aga	aaa	tat	gaa	gaa	acc	caa	1112
Leu	Ser	Phe	Glu	Leu	Ser	Glu	Phe	Arg	Arg	Lys	Tyr	Glu	Glu	Thr	Gln	
		260					265					270				
aaa	gaa	gtt	cac	aat	tta	aat	cag	ctg	ttg	tat	tca	caa	aga	agg	gca	1160
Lys	Glu	Val	His	Asn	Leu	Asn	Gln	Leu	Leu	Tyr	Ser	Gln	Arg	Arg	Ala	
		275				280					285					
gat	gtg	caa	cat	ctg	gaa	gat	gat	agg	cat	aaa	aca	gag	aag	ata	caa	1208
Asp	Val	Gln	His	Leu	Glu	Asp	Asp	Arg	His	Lys	Thr	Glu	Lys	Ile	Gln	
					295					300					305	
aaa	ctc	agg	gaa	gag	aat	gat	att	gct	agg	gga	aaa	ctt	gaa	gaa	gag	1256
Lys	Leu	Arg	Glu	Glu	Asn	Asp	Ile	Ala	Arg	Gly	Lys	Leu	Glu	Glu	Glu	
				310					315					320		
aag	aag	aga	tcc	gaa	gag	ctc	tta	tct	cag	gtc	cag	ttt	ctt	tac	aca	1304
Lys	Lys	Arg	Ser	Glu	Glu	Leu	Leu	Ser	Gln	Val	Gln	Phe	Leu	Tyr	Thr	
			325					330					335			
tct	ctg	cta	aag	cag	caa	gaa	gaa	caa	aca	agg	gta	gct	ctg	ttg	gaa	1352
Ser	Leu	Leu	Lys	Gln	Gln	Glu	Glu	Gln	Thr	Arg	Val	Ala	Leu	Leu	Glu	

340	345	350	
caa cag atg cag gca tgt act tta gac ttt gaa aat gaa aaa ctc gac			1400
Gln Gln Met Gln Ala Cys Thr Leu Asp Phe Glu Asn Glu Lys Leu Asp			
355	360	365	
cgt caa cat gtg cag cat caa ttg cat gta att ctt aag gag ctc cga			1448
Arg Gln His Val Gln His Gln Leu His Val Ile Leu Lys Glu Leu Arg			
370	375	380	385
aaa gca aga aat caa ata aca cag ttg gaa tcc ttg aaa cag ctt cat			1496
Lys Ala Arg Asn Gln Ile Thr Gln Leu Glu Ser Leu Lys Gln Leu His			
	390	395	400
gag ttt gcc atc aca gag cca tta gtc act ttc caa gga gag act gaa			1544
Glu Phe Ala Ile Thr Glu Pro Leu Val Thr Phe Gln Gly Glu Thr Glu			
	405	410	415
aac aga gaa aaa gtt gcc gcc tca cca aaa agt ccc act gct gca ctc			1592
Asn Arg Glu Lys Val Ala Ala Ser Pro Lys Ser Pro Thr Ala Ala Leu			
	420	425	430
aat gaa agc ctg gtg gaa tgt ccc aag tgc aat ata cag tat cca gcc			1640
Asn Glu Ser Leu Val Glu Cys Pro Lys Cys Asn Ile Gln Tyr Pro Ala			
	435	440	445
act gag cat cgc gat ctg ctt gtc cat gtg gaa tac tgt tca aag tag			1688
Thr Glu His Arg Asp Leu Leu Val His Val Glu Tyr Cys Ser Lys *			
450	455	460	
caaaataagt atttgttttg atattaaaag attcaatact gtattttctg ttagcttggtg			1748
ggcatttttga attatatatt tcacattttg cataaaactg cctatctacc tttgacactc			1808
cagcatgcta gtgaatcatg tatcttttag gctgctgtgc atttctcttg gcagtgatac			1868
ctccctgaca tggttcatca tcaggctgca atgacagaat gtggtgagca gcgtctactg			1928
agactactaa cattttgcac tgtcaaaata cttggtgagg aaaagatagc tcagggttatt			1988
gctaattgggt taatgcacca gcaagcaaaa tattttatgt tttgggggtt tgaaaaatca			2048
aagataatta accaaggatc ttaactgtgt tgcattttt tatccaagca cttagaaaac			2108
ctacaatcct aattttgatg tccattgtta agaggtggtg atagatacta tttttttttt			2168
catattgtat agcggttatt agaaaagttg gggattttct tgatctttat tgctgcttac			2228
cattgaaact taaccagct gtgttcccca actctgttct gcgcacgaaa cagtatctgt			2288
ttgaggcata atcttaagtg gccacacaca atgttttctc ttatgttatc tggcagtaac			2348
tgtaacttga attacattag cacattctgc ttagctaaaa ttgttaaaat aaactttaat			2408
aaacccatgt agccctctca tttgattgac agtatttttag ttatttttgg cattcttaaa			2468
gctgggcaat gtaatgatca gatctttgtt tgtctgaaca ggtattttta tacatgcttt			2528
ttgtaaaacca aaaactttta aatttcttca gggttttctaa catgcttacc actgggctac			2588
tgtaaatgag aaaagaataa aattatttta tgtttt			2624
 <210> 2			
<211> 464			
<212> PRT			
<213> Homo sapiens			
 <400> 2			
Met Ser Ser Arg Ser Thr Lys Asp Leu Ile Lys Ser Lys Trp Gly Ser			
1 5 10 15			
Lys Pro Ser Asn Ser Lys Ser Glu Thr Thr Leu Glu Lys Leu Lys Gly			
20 25 30			
Glu Ile Ala His Leu Lys Thr Ser Val Asp Glu Ile Thr Ser Gly Lys			
35 40 45			
Gly Lys Leu Thr Asp Lys Glu Arg His Arg Leu Leu Glu Lys Ile Arg			
50 55 60			

Val	Leu	Glu	Ala	Glu	Lys	Glu	Lys	Asn	Ala	Tyr	Gln	Leu	Thr	Glu	Lys
65					70					75					80
Asp	Lys	Glu	Ile	Gln	Arg	Leu	Arg	Asp	Gln	Leu	Lys	Ala	Arg	Tyr	Ser
				85					90					95	
Thr	Thr	Ala	Leu	Leu	Glu	Gln	Leu	Glu	Glu	Thr	Thr	Arg	Glu	Gly	Glu
			100					105					110		
Arg	Arg	Glu	Gln	Val	Leu	Lys	Ala	Leu	Ser	Glu	Glu	Lys	Asp	Val	Leu
		115					120					125			
Lys	Gln	Gln	Leu	Ser	Ala	Ala	Thr	Ser	Arg	Ile	Ala	Glu	Leu	Glu	Ser
	130					135					140				
Lys	Thr	Asn	Thr	Leu	Arg	Leu	Ser	Gln	Thr	Val	Ala	Pro	Asn	Cys	Phe
145					150					155					160
Asn	Ser	Ser	Ile	Asn	Asn	Ile	His	Glu	Met	Glu	Ile	Gln	Leu	Lys	Asp
				165				170						175	
Ala	Leu	Glu	Lys	Asn	Gln	Gln	Trp	Leu	Val	Tyr	Asp	Gln	Gln	Arg	Glu
			180					185					190		
Val	Tyr	Val	Lys	Gly	Leu	Leu	Ala	Lys	Ile	Phe	Glu	Leu	Glu	Lys	Lys
	195						200					205			
Thr	Glu	Thr	Ala	Ala	His	Ser	Leu	Pro	Gln	Gln	Thr	Lys	Lys	Pro	Glu
	210					215					220				
Ser	Glu	Gly	Tyr	Leu	Gln	Glu	Glu	Lys	Gln	Lys	Cys	Tyr	Asn	Asp	Leu
225					230					235					240
Leu	Ala	Ser	Ala	Lys	Lys	Asp	Leu	Glu	Val	Glu	Arg	Gln	Thr	Ile	Thr
				245				250						255	
Gln	Leu	Ser	Phe	Glu	Leu	Ser	Glu	Phe	Arg	Arg	Lys	Tyr	Glu	Glu	Thr
			260					265					270		
Gln	Lys	Glu	Val	His	Asn	Leu	Asn	Gln	Leu	Leu	Tyr	Ser	Gln	Arg	Arg
	275						280					285			
Ala	Asp	Val	Gln	His	Leu	Glu	Asp	Asp	Arg	His	Lys	Thr	Glu	Lys	Ile
	290					295					300				
Gln	Lys	Leu	Arg	Glu	Glu	Asn	Asp	Ile	Ala	Arg	Gly	Lys	Leu	Glu	Glu
305					310				315						320
Glu	Lys	Lys	Arg	Ser	Glu	Glu	Leu	Leu	Ser	Gln	Val	Gln	Phe	Leu	Tyr
				325				330					335		
Thr	Ser	Leu	Leu	Lys	Gln	Gln	Glu	Glu	Gln	Thr	Arg	Val	Ala	Leu	Leu
		340						345				350			
Glu	Gln	Gln	Met	Gln	Ala	Cys	Thr	Leu	Asp	Phe	Glu	Asn	Glu	Lys	Leu
	355						360					365			
Asp	Arg	Gln	His	Val	Gln	His	Gln	Leu	His	Val	Ile	Leu	Lys	Glu	Leu
	370					375					380				
Arg	Lys	Ala	Arg	Asn	Gln	Ile	Thr	Gln	Leu	Glu	Ser	Leu	Lys	Gln	Leu
385					390					395					400
His	Glu	Phe	Ala	Ile	Thr	Glu	Pro	Leu	Val	Thr	Phe	Gln	Gly	Glu	Thr
				405				410					415		
Glu	Asn	Arg	Glu	Lys	Val	Ala	Ala	Ser	Pro	Lys	Ser	Pro	Thr	Ala	Ala
		420						425				430			
Leu	Asn	Glu	Ser	Leu	Val	Glu	Cys	Pro	Lys	Cys	Asn	Ile	Gln	Tyr	Pro
	435						440					445			
Ala	Thr	Glu	His	Arg	Asp	Leu	Leu	Val	His	Val	Glu	Tyr	Cys	Ser	Lys
	450					455					460				

<210> 3

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> An artificially synthesized primer sequence for
RT-PCR

<400> 3
 acaacagcct caagatcatc ag 22

 <210> 4
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized primer sequence for
 RT-PCR

 <400> 4
 ggtccaccac tgacacgttg 20

 <210> 5
 <211> 23
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized primer sequence for
 RT-PCR

 <400> 5
 agagatccga agagctctta tct 23

 <210> 6
 <211> 21
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized primer sequence for
 RT-PCR

 <400> 6
 gatgctcagt ggctggatac t 21

 <210> 7
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized primer sequence for
 RT-PCR

 <400> 7
 cgaaagcttc agagatgtct tcca 24

 <210> 8
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized primer sequence for
 RT-PCR

 <400> 8

aatggatccc tttgaacagt attccac	27
<210> 9	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> An artificially synthesized primer sequence for RT-PCR	
<400> 9	
atagaattca tgtcttccag aagtac	26
<210> 10	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> An artificially synthesized primer sequence for RT-PCR	
<400> 10	
tatctcgagc tttgaacagt at	22
<210> 11	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> An artificially synthesized primer sequence for siRNA vector	
<400> 11	
tggtagccaa gtgcaggtta ta	22
<210> 12	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> An artificially synthesized primer sequence for siRNA vector	
<400> 12	
ccaaagggtt tctgcagttt ca	22
<210> 13	
<211> 30	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> An artificially synthesized primer sequence for siRNA vector	
<400> 13	
tgcggatcca gagcagattg tactgagagt	30

<210> 14
 <211> 29
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized primer sequence for
 siRNA vector

 <400> 14
 ctctatctcg agtgaggcgg aaagaacca 29

 <210> 15
 <211> 47
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized primer sequence for
 siRNA vector

 <400> 15
 tttaagcttg aagaccattt ttggaaaaaa aaaaaaaaaa aaaaaaac 47

 <210> 16
 <211> 34
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized primer sequence for
 siRNA vector

 <400> 16
 tttaagcttg aagacatggg aaagagtggc ctca 34

 <210> 17
 <211> 51
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized oligonucleotide
 sequence for siRNA

 <400> 17
 caccgaagca gcacgacttc ttcttcaaga gagaagaagt cgtgctgctt c 51

 <210> 18
 <211> 51
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized oligonucleotide
 sequence for siRNA

 <400> 18
 aaaagaagca gcacgacttc ttctctcttg aagaagaagt cgtgctgctt c 51

<210> 19
 <211> 51
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized oligonucleotide
 sequence for siRNA

 <400> 19
 tcccggagag actgaaaaca gagttcaaga gactctgttt tcagtctctc c 51

 <210> 20
 <211> 51
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized oligonucleotide
 sequence for siRNA

 <400> 20
 aaaaggagag actgaaaaca gagtctcttg aactctgttt tcagtctctc c 51

 <210> 21
 <211> 19
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized oligonucleotide
 sequence for siRNA target

 <400> 21
 ggagagactg aaaacagag 19

 <210> 22
 <211> 47
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized sequence of hairpin
 loop structure of siRNA

 <400> 22
 ggagagactg aaaacagagt tcaagagact ctgttttcag tctctcc 47

 <210> 23
 <211> 5089
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> An artificially synthesized vector sequence

 <221> misc_feature
 <222> (489)...(492)
 <223> n = GAP

<400> 23

gacggatcgg	gagatctccc	gatccccat	ggtgcactct	cagtacaatc	tgctctggat	60
ccactagtaa	cggccgccag	tgtgctggaa	ttcggcttgg	tagccaagtg	caggttatag	120
ggagctgaag	ggaagggggt	cacagtaggt	ggcatcggtc	ctttctgact	gcccgccecc	180
cgcatgccgt	cccgcgatat	tgagctccga	acctctcgcc	ctgccgcgcg	cggtgctccg	240
tcgccgcgcg	gccgccatgg	aattcgaaag	ctgacgtcat	caaccgcgtc	caaggaatcg	300
cgggcccagt	gtcactaggg	gggaacaccc	agcgcgcgtg	cgccctggca	ggaagatggc	360
tgtgagggac	aggggagtg	cgccctgcaa	tatttgcatg	tcgctatgtg	ttctgggaaa	420
tcaccataaa	cgtgaaatgt	ctttggattt	gggaatctta	taagttctgt	atgagaccac	480
tctttcccn	nntttttggg	aaaaaaaaa	aaaaaaaaa	aacgaaaccg	ggccggggcg	540
ggtggttcac	gcctataatc	ccagcacttt	gggaggccga	ggcggggcga	tcacaaggtc	600
aggaggtcga	gaccatccag	gctaacacgg	tgaaaccccc	ccccatctct	actaaaaaaa	660
aaaaatacaa	aaaattagcc	attagccggg	cgtgggtggc	ggcgcctata	atcccagcta	720
cttgggaggg	tgaagcagaa	tggcgtgaac	ccgggagggc	gacgttgca	tgagccgaga	780
tcgcgcgcg	tgcatccag	cctgggcgac	agagcgagtc	tcaaaaaaaa	aaccgagtg	840
aatgtgaaaa	gctccgtgaa	actgcagaaa	cccaagccga	attctgcaga	tatccatcac	900
actggcggcc	gctcgagtga	ggcggaaaga	accagctggg	gctctagggg	gtatccccac	960
gcgccttgta	gcggcgcat	aagcgcggcg	ggtgtggtgg	ttacgcgcag	cgtgaccgct	1020
acacttgcca	gcgccttagc	gcccgtcctt	ttcgctttct	tcccttcctt	tctcgccacg	1080
ttcgccggct	ttccccgtca	agctctaaat	cgggggctcc	ctttaggggt	ccgatttagt	1140
gctttacggc	acctcgaccc	caaaaaactt	gattaggggt	atggttcacg	tagtgggcca	1200
tcgccctgat	agacggtttt	tcgccctttg	acgttgagtg	ccacgttctt	taatagtga	1260
ctcttgttcc	aaactggaac	aacactcaac	cctatctcgg	tctattcttt	tgatttataa	1320
gggattttgc	cgatttcggc	ctattgggta	aaaaatgagc	tgatttaaca	aaaatttaac	1380
gcgaattaat	tctgtggaat	gtgtgtcagt	taggggtgtg	aaagtcccca	ggctccccag	1440
caggcagaag	tatgcaaagc	atgcatctca	attagtcagc	aaccaggtgt	ggaaagtccc	1500
caggctcccc	agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	gcaaccatag	1560
tccccgccct	aactccgccc	atcccccccc	taactccgcc	cagttccgcc	cattctccgc	1620
cccatggctg	actaattttt	tttatttatg	cagaggccga	ggccgcctct	gcctctgagc	1680
tattccagaa	gtagtggaga	ggcttttttg	gaggccatag	cttttgcaaa	aagctcccgg	1740
gagcttgat	atccattttc	ggatctgatt	aagagacagg	atgaggatcg	tttcgcatga	1800
ttgaacaaga	tggattgcac	gcaggttctc	cggccgcttg	ggtggagagg	ctattcggct	1860
atgactgggc	acaacagaca	atcggtgct	ctgatgccgc	cgtgttcggg	ctgtcagcgc	1920
aggggcgccc	ggttcttttt	gtcaagaccg	acctgtccgg	tgccctgaat	gaactgcagg	1980
acgaggcagc	gcggctatcg	tggttgccca	cgacgggcgt	tccttgcgca	gctgtgctcg	2040
acgttgtcac	tgaagcggga	agggactggc	tgctattggg	cgaagtgccg	gggcaggatc	2100
tcctgtcatc	tcaccttgct	cctgccgaga	aagtatccat	catggctgat	gcaatgcggc	2160
ggctgcatac	gcttgatccg	gctacctgcc	cattcgacca	ccaagcgaaa	catcgcatcg	2220
agcagcacg	tactcgatg	gaagccggtc	ttgtcgatca	ggatgatctg	gacgaagacg	2280
atcaggggct	cgccccagcc	gaactgttcg	ccaggctcaa	ggcgcgcgat	cccgcggcgg	2340
aggatctcgt	cgtgacccat	ggcgatgcct	gcttgccgaa	tatcatggtg	gaaaatggcc	2400
gcttttctgg	attcatcgac	tgtggccggc	tgggtgtggc	ggaccgctat	caggacatag	2460
cgttggttac	ccgtgatatt	gctgaagagc	ttggcggcga	atgggctgac	cgcttctctg	2520
tgctttacgg	tatcgccgct	cccgatccgc	agcgcacgc	cttctatcgc	cttcttgacg	2580
agttcttctg	agcgggactc	tggggttcga	aatgaccgac	caagcgacgc	ccaacctgcc	2640
atcacgagat	ttcgattcca	ccgccgcctt	ctatgaaagg	ttgggcttcg	gaatcgtttt	2700
ccgggacgcc	ggctggatga	tcctccagcg	cggggatctc	atgctggagt	tcctcgccca	2760
ccccaaactg	tttattgcag	cttataatgg	ttacaaataa	agcaatagca	tcacaaattt	2820
cacaaataaa	gcattttttt	cactgcattc	tagttgtggt	ttgtccaaac	tcataaatgt	2880
atcttatcat	gtctgtatac	cgtcgacctc	tagctagagc	ttggcgtaat	catggtcata	2940
gctgtttcct	gtgtgaaatt	gttatccgct	cacaattcca	cacaacatac	gagccggaag	3000
cataaagtgt	aaagcctggg	gtgcctaatt	agtgaagctaa	ctcacattaa	ttgcgttgcg	3060
ctcactgccc	gctttccagt	cgggaaacct	gtcgtgccag	ctgcattaat	gaatcggcca	3120
acgcgcgggg	agaggcggtt	tgcgtatttg	gcgctcttcc	gcttctctgc	tactgactc	3180
gctgcgctcg	gtcgttcggc	tgccggcgagc	ggtatcagct	cactcaaagg	cggtaatatc	3240
ggtatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaa	gccagcaaaa	3300
ggccaggaac	cgtaaaaaag	ccgcgttgct	ggcggttttc	cataggctcc	gcccccttga	3360
cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aaccgcagag	gactataaag	3420
ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttcgga	ccctgccgct	3480
taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	3540
ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgcacgaacc	3600

ccccgttcag	cccgaccgct	gcgccttata	cggtaactat	cgtcttgagt	ccaacccggt	3660
aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	3720
tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	3780
agtatttggg	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	3840
ttgatccggc	aaacaaacca	ccgctggtag	cggttttttt	gtttgcaagc	agcagattac	3900
gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggg	ctgacgctca	3960
gtggaacgaa	aactcacggt	aagggatttt	ggtcagtaga	ttatcaaaaa	ggatcttcac	4020
ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	4080
ttggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	4140
tcgttcatcc	atagttgcct	gactccccgt	cgtgtagata	actacgatac	gggaggggct	4200
accatctggc	cccagtgctg	caatgatacc	gcgagaccca	cgtccaccgg	ctccagattt	4260
atcagcaata	aaccagccag	ccggaagggc	cgagcgcaga	agtggtcctg	caactttatc	4320
cgctcccatc	cagtctatta	attgttgccg	ggaagctaga	gtaagtagtt	cgccagttaa	4380
tagtttgccg	aacgttggtg	ccattgctac	aggcatcggt	gtgtcacgct	cgtcgtttgg	4440
tatgggttca	ttcagctccg	gttcccaacg	atcaaggcga	gttacatgat	cccccatggt	4500
gtgcaaaaaa	gcggttagct	ccttcgggtc	tccgatcggt	gtcagaagta	agttggccgc	4560
agtgttatca	ctcatgggta	tggcagcact	gcataattct	cttactgtca	tgccatccgt	4620
aagatgcttt	tctgtgactg	gtgagtactc	aaccaagtca	ttctgagaat	agtgtatgcg	4680
gcgaccgagt	tgctcttgcc	cggcgtcaat	acgggataat	accgcgccac	atagcagaac	4740
tttaaaagtg	ctcatcattg	gaaaacgttc	ttcggggcga	aaactctcaa	ggatctttacc	4800
gctgttgaga	tccagttcga	tgtaacccac	tcgtgcaccc	aactgatctt	cagcatcttt	4860
tactttcacc	agcgtttctg	ggtgagcaaa	aacaggaagg	caaaatgccg	caaaaaaggg	4920
aataagggcg	acacggaaat	gttgaatact	catactcttc	ctttttcaat	attattgaag	4980
catttatcag	ggttattgtc	tcatgagcgg	atacatatct	gaatgtatct	agaaaaataa	5040
acaaataggg	gttccgcgca	catttccccg	aaaagtgcc	cctgacgtc		5089